.OTECHNOLOGY
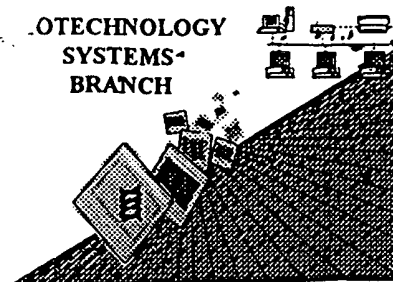SYSTEMS·
BRANCH

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information
Center (STIC) detected errors when processing the following computer readable
form:

Application Serial Number: _09/941947_

Source: _OIPE_

Date Processed by STIC: _09│18│2001_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE**
   APPLICANT, WITH A NOTICE TO COMPLY or,
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A**
   NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov **or phone 703-306-4119 (R. Wax)**
PATENTIN 3.0 e-mail help: patin3help@uspto.gov **or phone 703-306-4119 (R. Wax)**

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program
employing a logical and intuitive user-interface to check whether a sequence listing is in
compliance with format and content rules. Checker Version 3.0 works for sequence listings
generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old
rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual
Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-
compliant. Checker allows public users to check sequence listings in Computer Readable form
(CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence
listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker·

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/941947

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOF1

1_____Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2_____Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3_____Misaligned Amino
Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4_____Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5_____Variable Length

Sequence(s)____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6_____PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7_____Skipped Sequences
(OLD RULES)

Sequence(s)_____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8_____Skipped Sequences
(NEW RULES)

Sequence(s)_____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9_____Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10_____Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11___✓__Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12_____PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13___.___Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

RAW SEQUENCE LISTING   DATE: 09/18/2001
PATENT APPLICATION: US/09/941,947   TIME: 10:40:16

Input Set : A:\CL1903 US NA Seq Listing.txt
Output Set: N:\CRF3\09182001\I941947.raw

```
  5 <110> APPLICANT: Brzostowicz, Patricia C.
  6        Cheng, Qiong
  7        DiCosimo, Deana J.
  8        Koffas, Mattheos
  9        Miller, Edward S. Jr.
 10        Odom, J. Martin
 11        Picataggio, Steve
 12        Rouviere, Pierre E.
 16 <120> TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
 20 <130> FILE REFERENCE: CL1903 US NA
 23 <140> CURRENT APPLICATION NUMBER: US/09/941,947
 23 <141> CURRENT FILING DATE: 2001-08-29
 23 <150> PRIOR APPLICATION NUMBER: 60/229,907
 24 <151> PRIOR FILING DATE: 2000-09-01
 26 <150> PRIOR APPLICATION NUMBER: 60/229,858
 27 <151> PRIOR FILING DATE: 2000-09-01
 29 <160> NUMBER OF SEQ ID NOS: 60
 33 <170> SOFTWARE: Microsoft Office 97
 37 <210> SEQ ID NO: 1
 39 <211> LENGTH: 1311
 41 <212> TYPE: DNA
 43 <213> ORGANISM: Methylomonas 16a
 47 <400> SEQUENCE: 1
 48 gatgtggtca catggcccta tcacttaacg gctgatattc gattttgtca ttggtttttt      60
 50 cttaacttta acttctacac gctcatgaac aaacctaaaa aagttgcaat actgacagca     120
 52 ggcggcttgg cgccttgttt gaattccgca atcggtagtt tgatcgaacg ttataccgaa     180
 54 atcgatccta gcatagaaat catttgctat cgcggcggtt ataaaggcct gttgctgggc     240
 56 gattcttatc cagtaacggc cgaagtgcgt aaaaaggcgg gtgttctgca acgttttggc     300
 58 ggttctgtga tcggcaacag ccgcgtcaaa ttgaccaatg tcaaagactg cgtgaaacgc     360
 60 ggtttggtca agagggtga agatccgcaa aaagtcgcgg ctgatcaatt ggttaaggat     420
 62 ggtgtcgata ttctgcacac catcggcggc gatgatacca atacggcagc agcggatttg     480
 64 gcagcattcc tggccagaaa taattacgga ctgaccgtca ttggtttacc taaaaccgtc     540
 66 gataacgacg tatttccgat caagcaatca ctaggtgctt ggactgccgc cgagcaaggc     600
 68 gcgcgttatt tcatgaacgt ggtggccgaa acaacgcca acccacgcat gctgatcgta     660
 70 cacgaagtga tgggccgtaa ctgcggctgg ctgaccgctg caaccgcgca ggaatatcgc     720
 72 aaattactgg accgtgccga gtggttgccg gaattgggtt tgactcgtga atcttatgaa     780
 74 gtgcacgcgg tattcgttcc ggaaatggcg atcgacctgg aagccgaagc caagcgcctg     840
 76 cgcgaagtga tggacaaagt cgattgcgtc aacatcttcg tttccgaagg tgccggcgtc     900
 78 gaagctatcg tcgcggaaat gcaggccaaa ggccaggaag tgccgcgcga tgcgttcggc     960
 80 cacatcaaac tggatgcggt caaccctggt aaatggttcg gcgagcaatt cgcgcagatg    1020
 82 ataggcgcgg aaaaaaccct ggtacaaaaa tcgggatact tcgcccgtgc ttctgcttcc    1080
 84 aacgttgacg acatgcgttt gatcaaatcg tgcgccgact ggcggtcga gtgcgcgttc    1140
 86 cgccgcgagt ctggcgtgat cggtcacgac gaagacaacg gcaacgtgtt gcgtgcgatc    1200
 88 gagtttccgc gcatcaaggg cggcaaaccg ttcaatatcg acaccgactg gttcaatagc    1260
 90 atgttgagcg aaatcggcca gcctaaaggc ggtaaagtcg aagtcagcca c             1311
 93 <210> SEQ ID NO: 2
 95 <211> LENGTH: 437
```

RAW SEQUENCE LISTING   DATE: 09/18/2001
PATENT APPLICATION: US/09/941,947   TIME: 10:40:16

Input Set : A:\CL1903 US NA Seq Listing.txt
Output Set: N:\CRF3\09182001\I941947.raw

97 <212> TYPE: PRT
99 <213> ORGANISM: Methylomonas 16a
103 <400> SEQUENCE: 2
105 Asp Val Val Thr Trp Pro Tyr His Leu Thr Ala Asp Ile Arg Phe Cys
106 1   5   10   15
109 His Trp Phe Phe Leu Asn Phe Asn Phe Tyr Thr Leu Met Asn Lys Pro
110   20   25   30
113 Lys Lys Val Ala Ile Leu Thr Ala Gly Gly Leu Ala Pro Cys Leu Asn
114   35   40   45
117 Ser Ala Ile Gly Ser Leu Ile Glu Arg Tyr Thr Glu Ile Asp Pro Ser
118   50   55   60
121 Ile Glu Ile Ile Cys Tyr Arg Gly Gly Tyr Lys Gly Leu Leu Leu Gly
122 65   70   75   80
125 Asp Ser Tyr Pro Val Thr Ala Glu Val Arg Lys Lys Ala Gly Val Leu
126   85   90   95
129 Gln Arg Phe Gly Gly Ser Val Ile Gly Asn Ser Arg Val Lys Leu Thr
130   100   105   110
133 Asn Val Lys Asp Cys Val Lys Arg Gly Leu Val Lys Glu Gly Glu Asp
134   115   120   125
137 Pro Gln Lys Val Ala Ala Asp Gln Leu Val Lys Asp Gly Val Asp Ile
138   130   135   140
141 Leu His Thr Ile Gly Gly Asp Asp Thr Asn Thr Ala Ala Ala Asp Leu
142 145   150   155   160
145 Ala Ala Phe Leu Ala Arg Asn Asn Tyr Gly Leu Thr Val Ile Gly Leu
146   165   170   175
149 Pro Lys Thr Val Asp Asn Asp Val Phe Pro Ile Lys Gln Ser Leu Gly
150   180   185   190
153 Ala Trp Thr Ala Ala Glu Gln Gly Ala Arg Tyr Phe Met Asn Val Val
154   195   200   205
157 Ala Glu Asn Asn Ala Asn Pro Arg Met Leu Ile Val His Glu Val Met
158   210   215   220
161 Gly Arg Asn Cys Gly Trp Leu Thr Ala Ala Thr Ala Gln Glu Tyr Arg
162 225   230   235   240
165 Lys Leu Leu Asp Arg Ala Glu Trp Leu Pro Glu Leu Gly Leu Thr Arg
166   245   250   255
169 Glu Ser Tyr Glu Val His Ala Val Phe Val Pro Glu Met Ala Ile Asp
170   260   265   270
173 Leu Glu Ala Glu Ala Lys Arg Leu Arg Glu Val Met Asp Lys Val Asp
174   275   280   285
177 Cys Val Asn Ile Phe Val Ser Glu Gly Ala Gly Val Glu Ala Ile Val
178   290   295   300
181 Ala Glu Met Gln Ala Lys Gly Gln Glu Val Pro Arg Asp Ala Phe Gly
182 305   310   315   320
185 His Ile Lys Leu Asp Ala Val Asn Pro Gly Lys Trp Phe Gly Glu Gln
186   325   330   335
189 Phe Ala Gln Met Ile Gly Ala Glu Lys Thr Leu Val Gln Lys Ser Gly
190   340   345   350
193 Tyr Phe Ala Arg Ala Ser Ala Ser Asn Val Asp Asp Met Arg Leu Ile
194   355   360   365

```
197 Lys Ser Cys Ala Asp Leu Ala Val Glu Cys Ala Phe Arg Arg Glu Ser
198         370             375             380
201 Gly Val Ile Gly His Asp Glu Asp Asn Gly Asn Val Leu Arg Ala Ile
202 385             390             395             400
205 Glu Phe Pro Arg Ile Lys Gly Gly Lys Pro Phe Asn Ile Asp Thr Asp
206             405             410             415
209 Trp Phe Asn Ser Met Leu Ser Glu Ile Gly Gln Pro Lys Gly Gly Lys
210         420             425             430
213 Val Glu Val Ser His
214         435
217 <210> SEQ ID NO: 3
219 <211> LENGTH: 636
221 <212> TYPE: DNA
223 <213> ORGANISM: Methylomonas 16a
227 <400> SEQUENCE: 3
228 gaaaatacta tgtccgtcac catcaaagaa gtcatgacca cctcgcccgt tatgccggtc   60
230 atggtcatca atcatctgga acatgccgtc cctctggctc gcgcgctagt cgacggtggc  120
232 ttgaaagttt tggagatcac attgcgcacg ccggtggcac tggaatgtat ccgacgtatc  180
234 aaagccgaag taccggacgc catcgtcggc gcgggcacca tcatcaaccc tcataccttg  240
236 tatcaagcga ttgacgccgg tgcggaattc atcgtcagcc ccggcatcac cgaaaatcta  300
238 ctcaacgaag cgctagcatc cggcgtgcct atcctgcccg gcgtcatcac acccagcgag  360
240 gtcatgcgtt tattggaaaa aggcatcaat gcgatgaaat ctttccggc tgaagccgcc  420
242 ggcggcatac cgatgctgaa atcccttggc ggcccttgc cgcaagtcac cttctgtccg  480
244 accggcggcg tcaatcccaa aaacgcgccc gaatatctgg cattgaaaaa tgtcgcctgc  540
246 gtcggcggct cctggatggc gccggccgat ctggtagatg ccgaagactg ggcggaaatc  600
248 acgcggcggg cgagcgaggc cgcggcattg aaaaaa                             636
251 <210> SEQ ID NO: 4
253 <211> LENGTH: 212
255 <212> TYPE: PRT
257 <213> ORGANISM: Methylomonas 16a
261 <400> SEQUENCE: 4
263 Glu Asn Thr Met Ser Val Thr Ile Lys Glu Val Met Thr Thr Ser Pro
264 1               5               10              15
267 Val Met Pro Val Met Val Ile Asn His Leu Glu His Ala Val Pro Leu
268         20              25              30
271 Ala Arg Ala Leu Val Asp Gly Gly Leu Lys Val Leu Glu Ile Thr Leu
272         35              40              45
275 Arg Thr Pro Val Ala Leu Glu Cys Ile Arg Arg Ile Lys Ala Glu Val
276     50              55              60
279 Pro Asp Ala Ile Val Gly Ala Gly Thr Ile Ile Asn Pro His Thr Leu
280 65              70              75              80
283 Tyr Gln Ala Ile Asp Ala Gly Ala Glu Phe Ile Val Ser Pro Gly Ile
284         85              90              95
287 Thr Glu Asn Leu Leu Asn Glu Ala Leu Ala Ser Gly Val Pro Ile Leu
288         100             105             110
291 Pro Gly Val Ile Thr Pro Ser Glu Val Met Arg Leu Leu Glu Lys Gly
292     115             120             125
295 Ile Asn Ala Met Lys Phe Phe Pro Ala Glu Ala Ala Gly Gly Ile Pro
296     130             135             140
```

```
299 Met Leu Lys Ser Leu Gly Gly Pro Leu Pro Gln Val Thr Phe Cys Pro
300 145             150             155             160
303 Thr Gly Gly Val Asn Pro Lys Asn Ala Pro Glu Tyr Leu Ala Leu Lys
304             165             170             175
307 Asn Val Ala Cys Val Gly Gly Ser Trp Met Ala Pro Ala Asp Leu Val
308             180             185             190
311 Asp Ala Glu Asp Trp Ala Glu Ile Thr Arg Arg Ala Ser Glu Ala Ala
312         195             200             205
315 Ala Leu Lys Lys
316      210
319 <210> SEQ ID NO: 5
321 <211> LENGTH: 1860
323 <212> TYPE: DNA
325 <213> ORGANISM: Methylomonas 16a
329 <400> SEQUENCE: 5
330 atgaaactga ccaccgacta tcccttgctt aaaaacatcc acacgccggc ggacatacgc    60
332 gcgctgtcca aggaccagct ccagcaactg gctgacgagg tgcgcggcta tctgacccac   120
334 acggtcagca tttccggcgg ccattttgcg gccggcctcg gcaccgtgga actgaccgtg   180
336 gccttgcatt atgtgttcaa tacccccgtc gatcagttgg tctgggacgt gggccatcag   240
338 gcctatccgc acaagattct gaccggtcgc aaggagcgca tgccgaccat cgcaccctg    300
340 ggcggggtgt cagcctttcc ggcgcgggac gagagcgaat acgatgcctt cggcgtcggc   360
342 cattccagca cctcgatcag cgcggcactg ggcatggcca ttgcgtcgca gctgcgcggc   420
344 gaagacaaga agatggtagc catcatcggc gacggttcca tcaccggcgg catggcctat   480
346 gaggcgatga atcatgccgg cgatgtgaat gccaacctgc tggtgatctt gaacgacaac   540
348 gatatgtcga tctcgccgcc ggtcggggcg atgaacaatt atctgaccaa ggtgttgtcg   600
350 agcaagtttt attcgtcggt gcgggaagag agcaagaaag ctctggccaa gatgccgtcg   660
352 gtgtgggaac tggcgcgcaa gaccgaggaa cacgtgaagg gcatgatcgt gcccggtacc   720
354 ttgttcgagg aattgggctt caattatttc ggcccgatcg acggccatga tgtcgagatg   780
356 ctggtgtcga ccctggaaaa tctgaaggat ttgaccgggc cggtattcct gcatgtggtg   840
358 accaagaagg gcaaaggcta tgcgccagcc gagaagacc cgttggccta ccatggcgtg   900
360 ccggctttcg atccgaccaa ggatttcctg cccaaggcgg cgccgtcgcc gcatccgacc   960
362 tataccgagg tgttcggccg ctggctgtgc gacatggcgg ctcaagacga gcgcttgctg  1020
364 ggcatcacgc cggcgatgcg cgaaggctct ggtttggtgg aattctcaca gaaatttccg  1080
366 aatcgctatt tcgatgtcgc catcgccgag cagcatgcgg tgaccttggc cgccggccag  1140
368 gcctgccagg gcgccaagcc ggtggtggcg atttattcca ccttcctgca acgcggttac  1200
370 gatcagttga tccacgacgt ggccttgcag aacttagata tgctctttgc actggatcgt  1260
372 gccggcttgg tcggcccgga tggaccgacc catgctggcg cctttgatta cagctacatg  1320
374 cgctgtattc cgaacatgct gatcatggct ccagccgacg agaacgagtg caggcagatg  1380
376 ctgaccaccg gcttccaaca ccatggcccg gcttcggtgc gctatccgcg cggcaaaggg  1440
378 cccggggcgg caatcgatcc gaccctgacc gcgctggaga tcggcaaggc cgaagtcaga  1500
380 caccacggca gccgcatcgc cattctggcc tggggcagca tggtcacgcc tgccgtcgaa  1560
382 gccggcaagc agctgggcgc gacggtggtg aacatgcgtt tcgtcaagcc gttcgatcaa  1620
384 gccttggtgc tggaattggc caggacgcac gatgtgttcg tcaccgtcga ggaaaacgtc  1680
386 atcgccggcg cgctggcag tgcgatcaac accttcctgc aggcgcagaa ggtgctgatg  1740
388 ccggtctgca acatcggcct gcccgaccgc ttcgtcgagc aaggtagtcg cgaggaattg  1800
390 ctcagcctgg tcggcctcga cagcaagggc atcctcgcca ccatcgaaca gttttgcgct  1860
393 <210> SEQ ID NO: 6
395 <211> LENGTH: 620
397 <212> TYPE: PRT
```

```
399 <213> ORGANISM: Methylomonas 16a
403 <400> SEQUENCE: 6
405 Met Lys Leu Thr Thr Asp Tyr Pro Leu Leu Lys Asn Ile His Thr Pro
406 1               5                   10                  15
409 Ala Asp Ile Arg Ala Leu Ser Lys Asp Gln Leu Gln Gln Leu Ala Asp
410          20                  25                  30
413 Glu Val Arg Gly Tyr Leu Thr His Thr Val Ser Ile Ser Gly Gly His
414          35                  40                  45
417 Phe Ala Ala Gly Leu Gly Thr Val Glu Leu Thr Val Ala Leu His Tyr
418      50                  55                  60
421 Val Phe Asn Thr Pro Val Asp Gln Leu Val Trp Asp Val Gly His Gln
422 65                  70                  75                  80
425 Ala Tyr Pro His Lys Ile Leu Thr Gly Arg Lys Glu Arg Met Pro Thr
426              85                  90                  95
429 Ile Arg Thr Leu Gly Gly Val Ser Ala Phe Pro Ala Arg Asp Glu Ser
430          100                 105                 110
433 Glu Tyr Asp Ala Phe Gly Val Gly His Ser Ser Thr Ser Ile Ser Ala
434          115                 120                 125
437 Ala Leu Gly Met Ala Ile Ala Ser Gln Leu Arg Gly Glu Asp Lys Lys
438      130                 135                 140
441 Met Val Ala Ile Ile Gly Asp Gly Ser Ile Thr Gly Gly Met Ala Tyr
442 145                 150                 155                 160
445 Glu Ala Met Asn His Ala Gly Asp Val Asn Ala Asn Leu Leu Val Ile
446              165                 170                 175
449 Leu Asn Asp Asn Asp Met Ser Ile Ser Pro Pro Val Gly Ala Met Asn
450          180                 185                 190
453 Asn Tyr Leu Thr Lys Val Leu Ser Ser Lys Phe Tyr Ser Ser Val Arg
454          195                 200                 205
457 Glu Glu Ser Lys Lys Ala Leu Ala Lys Met Pro Ser Val Trp Glu Leu
458      210                 215                 220
461 Ala Arg Lys Thr Glu Glu His Val Lys Gly Met Ile Val Pro Gly Thr
462 225                 230                 235                 240
465 Leu Phe Glu Glu Leu Gly Phe Asn Tyr Phe Gly Pro Ile Asp Gly His
466              245                 250                 255
469 Asp Val Glu Met Leu Val Ser Thr Leu Glu Asn Leu Lys Asp Leu Thr
470          260                 265                 270
473 Gly Pro Val Phe Leu His Val Val Thr Lys Lys Gly Lys Gly Tyr Ala
474          275                 280                 285
477 Pro Ala Glu Lys Asp Pro Leu Ala Tyr His Gly Val Pro Ala Phe Asp
478      290                 295                 300
481 Pro Thr Lys Asp Phe Leu Pro Lys Ala Ala Pro Ser Pro His Pro Thr
482 305                 310                 315                 320
485 Tyr Thr Glu Val Phe Gly Arg Trp Leu Cys Asp Met Ala Ala Gln Asp
486              325                 330                 335
489 Glu Arg Leu Leu Gly Ile Thr Pro Ala Met Arg Glu Gly Ser Gly Leu
490          340                 345                 350
493 Val Glu Phe Ser Gln Lys Phe Pro Asn Arg Tyr Phe Asp Val Ala Ile
494          355                 360                 365
497 Ala Glu Gln His Ala Val Thr Leu Ala Ala Gly Gln Ala Cys Gln Gly
```

09/941947

```
<210> SEQ ID NO 41
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence          Errored
<220> FEATURE:
<223> OTHER INFORMATION:
<400> SEQUENCE: 41
      agcagctagc ggaggaataa accatgagcg catttctc
38
```

```
<210> SEQ ID NO 42
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence          Errored
<220> FEATURE:
<223> OTHER INFORMATION: :
<400> SEQUENCE: 42
      gactagtcac gacctgctcg aacgac
26
```

Errored: When the ORGANISM field is "Artificial Sequence" an explanation is mandatory in field 223.

The type of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.

VERIFICATION SUMMARY   DATE: 09/18/2001
PATENT APPLICATION: US/09/941,947   TIME: 10:40:17

Input Set : A:\CL1903 US NA Seq Listing.txt
Output Set: N:\CRF3\09182001\I941947.raw

L:23 M:270 C: Current Application Number differs, Replaced Current Application No
L:23 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:3057 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:41
L:3061 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:3061 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:
L:3071 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:42
L:3075 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:3075 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION: